

*J. Agr. Sci. Tech. (2015) Vol. 17: 791-803*

## Prediction of Wheat Production Using Artificial Neural Networks and Investigating Indirect Factors Affecting It: Case Study in Canterbury Province, New Zealand

M. Safa<sup>1\*</sup>, S. Samarasinghe<sup>2</sup>, and M. Nejat<sup>3</sup>

### ABSTRACT

An artificial neural network (ANN) approach was used to model the wheat production. From an extensive data collection involving 40 farms in Canterbury, New Zealand, the average wheat production was estimated at 9.9 t ha<sup>-1</sup>. The final ANN model developed was capable of predicting wheat production under different conditions and farming systems using direct and indirect technical factors. After examining more than 140 different factors, 6 factors were selected as influential input into the model. The final ANN model can predict wheat production based on farm conditions (wheat area and irrigation frequency), machinery condition (tractor hp ha<sup>-1</sup> and number of passes of sprayer) and farm inputs (N and fungicides consumption) in Canterbury with an error margin of  $\pm 9\%$  ( $\pm 0.89$  t ha<sup>-1</sup>).

**Keywords:** Agricultural Production, Agricultural Systems, Modelling.

### INTRODUCTION

Wheat provides nearly 55% of the carbohydrate and 20% of the calories consumed (Breiman and Graur, 1995). Also, more than 40% of the world's grain is fed to livestock (Pimentel and Pimentel, 2008). Wheat is cultivated under a wide range of climatic conditions and it is consumed more than any other cereal grain (Singh *et al.*, 2007). Global production of bread wheat, in 2010, was 651 million tons (Mt), with an average yield of 3.0 t ha<sup>-1</sup> (FAO, 2012). The world's major bread wheat producing areas are in northern China, northern India, northern USA and the adjoining areas in Canada, and in Europe, Russia, Latin America, and Africa (Kole, 2006). Wheat covers around 25% of the total global area devoted to cereal crops (Singh *et al.*, 2007).

It is the staple food of nearly 35% of the world's population. Recent statistics show that the demand for wheat grows faster than for any other major crops. In the last few decades, the development of new seed varieties has increased the yield. However, in many areas, because of the use of old growing systems, yields have stayed at less than desired levels (Rosegrant *et al.*, 1995; Ozkan *et al.*, 2004).

The forecasted global demand for wheat, in 2020, varied between 840 and 1,050 Mt (Rosegrant *et al.*, 1995; Kronstad, 1998). To achieve this target, global production will need to increase by 1.6 to 2.6% annually from the present production. Increases in realized grain yield have provided about 90% of the growth in cereal production since 1950 (Mitchell *et al.*, 1997). For wheat, the global average grain yield must increase from the current 3 to 3.8 t ha<sup>-1</sup>.

<sup>1</sup>Department of Land Management and Systems, Lincoln University, New Zealand.

<sup>2</sup>Department of Environmental Management, Lincoln University, New Zealand.

<sup>3</sup>Department of Agricultural Science, Payame Noor University, Farahan Branch, Islamic Republic of Iran.

\*Corresponding author; e-mail: [majeed.safa@lincoln.ac.nz](mailto:majeed.safa@lincoln.ac.nz)



Several models and algorithms have been developed to predict yield of agricultural productions. Many authors have found linear correlation of yield with soil properties and environmental conditions (Sudduth *et al.*, 1997; Khakural *et al.*, 1999; Gemtos *et al.*, 2005). Many other studies have used linear methods especially multiple linear regressions to predict yields using soil properties (Sudduth *et al.*, 1997; Khakural *et al.*, 1999; Wendroth *et al.*, 1999). However, using nonlinear methods mainly artificial neural networks (ANNs) and fuzzy logics for yield prediction have become more common in recent studies (Salehi *et al.*, 1998; Kominakis *et al.*, 2002; Kaul *et al.*, 2005; Sharma *et al.*, 2007; Papageorgiou *et al.*, 2011; Papageorgiou *et al.*, 2013).

In some of these studies, the effects of different factors, mainly soil properties and farm inputs, were investigated on test plots (Uno *et al.*, 2005; Niska *et al.*, 2010) or information were collected from real farm conditions by survey. Most studies show nonlinear statistical methods can predict yield better than multiple linear regression models; however, Uno *et al.* (2005) couldn't find clear difference between the two methods.

Several uncontrolled factors influence agricultural production; therefore, even complex and mathematical models cannot give the accurate results (Papageorgiou *et al.*, 2011; Safa and Samarasinghe, 2011). Several ANNs were employed for setting target yields. Schultz *et al.* (2000) show the advantages of applying ANNs in agricultural studies. ANNs has been used to predict the yield of different agricultural production, for example, it was used to predict yield of corn (Uno *et al.*, 2005), wheat (Alvarez, 2009; Faramarzi *et al.*, 2010; Özdoğan, 2011), maize (Folberth *et al.*, 2012), soybean (Soares *et al.*, 2013), banana (Soares *et al.*, 2013) and milk production (Kominakis *et al.*, 2002; Sharma *et al.*, 2007). All available studies on arable production have concentrated on specific direct technical variables mainly soil properties and soil moisture and no article were found to develop

a model to predict crop yield based on a wide range of direct and indirect variables.

Several studies have used NN models for classification, prediction, and problem solving. NNs have been applied by researchers in a wide range of application areas, such as mathematics, engineering, medicine, economics, environment, and agriculture (Sözen, 2009). Many of researchers have applied neural networks in the modelling of various scenarios to solve different problems, in which no explicit formulations were available (Fang *et al.*, 2000). The main advantage of neural networks is that they are able to use prior information (historical underlying process data) to develop an accurate representation of the process or relationship of interest.

The benefits of using NN models are the simplicity of application and robustness in results. NN models have developed into a powerful approach that can approximate any nonlinear input-output mapping function to any degree of accuracy in an iterative manner. NN models have many attractive properties for modelling complex production systems, and some of these are: universal function approximation capability, resistance to noisy or missing data, accommodation of multiple non-linear variables with unknown interactions and good generalization ability (Hagan *et al.*, 2002).

In most studies, a feed-forward Multi-Layered Perception (MLP) paradigm consisting of one or more inputs, hidden layers, and output layer trained by back propagation (BP) is used. Due to its documented ability to model any function, MLP trained with BP is selected to develop apparatus, processes, and product prediction models (Hornik *et al.*, 1989; Heinzow and Tol 2003; Jebaraj and Iniyar, 2006). In the processing of inputs by the network, each neuron in the first layer (hidden layer) processes the weighted inputs through a transfer function to produce its output. The transfer functions may be a linear or a nonlinear function. There are several transfer functions, such as Logistic,

Hyperbolic-tangent, Gussian, and Sine. The output depends on the particular transfer function used. This output is then sent to the neurons in the next layer through weighted corrections and these neurons complete their outputs by processing the sum of weighted inputs through their transfer functions. When this layer is the output layer, neuron output is the predicted output.

Several methods of error estimation have been proposed. The mean square error (MSE) over all training patterns [(Equation (1))] is the most commonly used error indicator. *MSE* is very useful to compare different models; it shows the networks ability to predict the correct output. The *MSE* can be written as:

$$MSE = \frac{1}{2N} \sum_i^N (t_i - z_i)^2 \quad (1) \text{ Where, } t_i \text{ and } z_i$$

are the actual and the predicted output for the  $i^{\text{th}}$  training pattern, and  $N$  is the total number of training patterns (Samarasinghe, 2007). Root mean square error (RMSE) is another error estimation, which shows the error in the units of actual and predicted data.

It is better to solve any problem with the minimum number of variables. When the number of variables is notably high, especially when there are limited number of samples, data reduction is useful. Also, when some input variables correlate with one another, another problem that is called multicollinearity, will appear. Correlation between inputs reduces the chance of having a unique solution (Samarasinghe, 2007). The best common method for data reduction and removal of multicollinearity is principal component analysis (PCA). PCA is a useful method to select the most important uncorrelated variables. PCA uses the mean and variance of each input variable and the covariance between variables to create a covariance matrix (COV) (Samarasinghe, 2007) and transforms the COV to obtain independent components that are linear summations of the original inputs. The results allow to either pick individual original variables that are uncorrelated or

use independent components that are independent as inputs to the model.

## MATERIALS AND METHODS

This study was conducted over 35,300 hectares of irrigated and dry land wheat fields in Canterbury, New Zealand, which represents around 90% of the wheat area and wheat production in New Zealand (Statistics New Zealand, 2010). Canterbury is the largest region in New Zealand, with an area of 45,346 km<sup>2</sup> (Statistics New Zealand, 1999). There is a wide range of landscapes in Canterbury from sweeping coastlines and dry plains to rugged bush-covered mountain ranges. Canterbury soil comprises yellow-grey earths, and their associated stony soils, over a very thick layer of gravel covered by fine materials of variable thickness. These soils were appropriate for intensive cropping of cereals and fodder crops and high-density sheep grazing. The maximum daily average temperature in summer is between 20 and 23°C. Furthermore, the average annual rainfall in most areas is between 650-700 millimeters; however, the high mountains receive over 4,000 millimeters of rain annually (Statistics New Zealand, 1999, 2004).

In this study, a wide range of factors, around 140, including farmers' social status, age of tractors and equipment, power of tractors, number and size of paddocks were studied. Moreover, these indirect factors and wheat production were examined to design the model to predict wheat production.

For use in the ANN model, it was necessary to select a limited number of relevant and influential variables without any bias; therefore, all information was investigated carefully. There were around 140 original variables, each of which could be a potential input in the final model. The collected data indicated that some inputs can be dropped; for example, 39 farms were managed by owners;



therefore, farm ownership was eliminated from the process, or some of the operations were not commonly used; consequently, those machines or operations were eliminated from the analysis as well. Finally 63 columns of inputs and outputs were selected and saved in another spreadsheet. This information was used to draw the graphs and carry out statistical analysis using MS Excel and SPSS software, respectively.

A strong feature of this study was in selecting a few of the best variables from several inputs. In this study, for variable reduction, correlation and principal component analysis (PCA) were used. Initially, variables were selected on the basis of no significant correlation between them but high relationship to energy consumption. Out of the variables that had significant correlations to each other and to energy consumption, the one with the strongest correlation with energy consumption was selected and the other was removed. The selected variables were further reduced by using PCA to select the final most relevant set of variables. Specifically, the PCs were carefully studied to select the uncorrelated inputs based on their coefficients in each PC.

Finding appropriate variables is the first step of model creation. After processing original data and input reduction, six variables were selected: wheat area (ha), total tractor horsepower per hectare (hp ha<sup>-1</sup>), nitrogen consumption (kg), fungicide consumption (kg), irrigation frequency, and number of passes of sprayer. These variables were not correlated and they were selected by using Principal Component Analysis (PCA).

NNs can be successfully trained to describe the influence of direct and indirect factors on wheat production. The sample size used in this study was 40 farms. Initially, a sample of 30 farms was randomly selected for training, and the remaining sample of 10 farms was used for validation.

After several trials by using Peltarion Synapse software, a modular neural network with two hidden layers was selected. In the modular network structure, the model is characterized by a series of independent neural networks after the input layer, which operates on the inputs to achieve some subtasks of the task the network expects to perform. These subtasks are trained separately with different examples from the sample and their outputs are summed in the output layer. The structure of the model prepares the network to use simultaneously different model functions for the data.

The Quick Prop was used as the learning method; because, it was fast in reducing the error and finding the best model. Quick Prop implicitly uses the second derivative of error to adjust weights. In each iteration of Quick Prop, the update for the weights was regulated as follows:

$$w_{m+1} = w_m + \Delta w_m$$
$$\Delta w_m = \frac{d_m}{d_{m-1} - d_m} \Delta w_{m-1}$$
$$d_m = \sum_{n=1}^N \left[ \frac{\partial E}{\partial w_m} \right]_n$$

Where,  $\Delta w_m$  is the current weight increment,  $d_m$  is the average derivative of the error with respect to the weight for the current epoch  $m$ ; and  $\partial E / \partial w_m$  is the current error gradient for a particular input vector (Samarasinghe, 2007).

Preliminary trials indicated that two hidden layer networks gave better results than one hidden layer network. Different functions were tested and in the final model, hyperbolic tangent (tanh) function was selected for the first hidden layer and logistic function was applied for the second hidden layer.

It is important to note that various combinations of number of layers and number of neurons, and different functions, structures, and learning methods were examined to find the best model with minimum iteration. Number of neurons was optimized by using a genetic algorithm programme.

## RESULTS AND DISCUSSION

### Wheat Production and Its Correlation with other Factors

To develop a practical model, simple statistical analysis of yield and more understanding of link between yield and other parameters is necessary. Yield was one of the simplest parameters for comparing farms and farmers. For many farmers, quantity of yield was more important than quality and environmental impacts. They tried to produce more crops by improving techniques and machinery as well as increasing farm inputs. In this study, maximum and minimum yield ranged between 6 and 15 tons per hectare, and average yield was estimated around  $9.9 \text{ t ha}^{-1}$  (Table 1). Average yield in this study was  $1.4 \text{ t ha}^{-1}$  more than the national average yield, in 2007 (Statistics New Zealand, 2008).

The study showed that yield ( $\text{t ha}^{-1}$ ) on larger farms was less than on smaller farms and it was negatively correlated with the size of farm at  $r = -0.36$  (Figure 1). Nevertheless, there was no significant correlation between yield and either the size of wheat area or crop area.

A positive correlation between the wheat area (ha)/total farm (ha) and crop area (ha)/total farm (ha) indices and yield was found and the relationship for wheat of  $r = 0.53$  is shown in Figure 2. It was clear that in mixed farms, which produced crop and dairy together, the proportion of wheat area was less than in arable farms. This association confirmed that farmers focused on the more beneficial aspects of their farms. Therefore, the proportion of wheat area to the total farm would be more important than the size of wheat area in farm yield analysis. This may

be due to different reasons; for example, usually farmers produced crops they have more knowledge about and experience with. In other words, when farmers had experience on wheat (crop) production, the proportion of wheat areas on their farms increased. This would be correct even for arable farms; where farmers produced only crops, the yield and proportion of area dedicated to particular crops the farmers had experience with were higher than those for other crops. This will be an interesting subject for research in future studies.

There was a negative correlation between the yield and numbers of sheep and cows at  $r = -0.41$  and  $r = -0.38$ , respectively. These results indicated that the size of the farm and the proportion of crop area was a key factor to increase the yield in wheat production, and farmers who kept more livestock usually had lower wheat yields than farmers who concentrated on crop production.

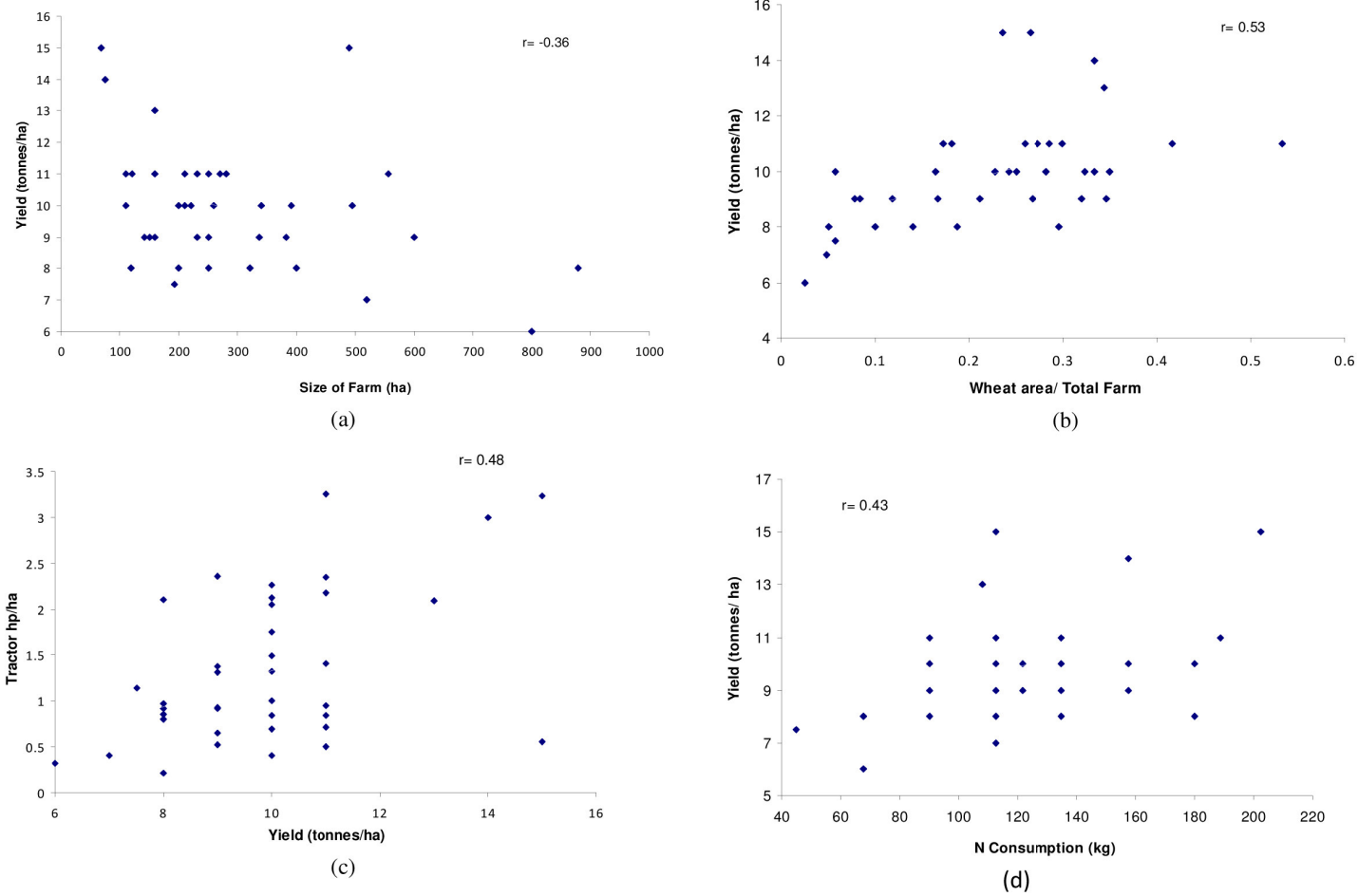
As shown in Figure 3, there was a positive significant relationship between yield and tractor power ( $\text{hp ha}^{-1}$ ) index, with  $r = 0.48$ . This indicated that as the power of tractors (mechanization) per hectare increased, the yield also increased, which explained why farmers have preferred to buy more powerful tractors and combines; however, the connection between the two factors should be investigated carefully.

The results of this study confirmed the farmer's opinion of the role of nitrogen in crop production. They believed that nitrogen (urea) was one of the most important factors to increase yield, revealed through the positive significant correlation between yield and nitrogen ( $r = 0.43$ ), as shown in Figure 4. It can be concluded that any plan to reduce nitrogen consumption, in current circumstances, would reduce wheat and other agricultural production.

As shown in Figure 5, fungicide

**Table 1.** Statistics of wheat production ( $\text{t ha}^{-1}$ ).

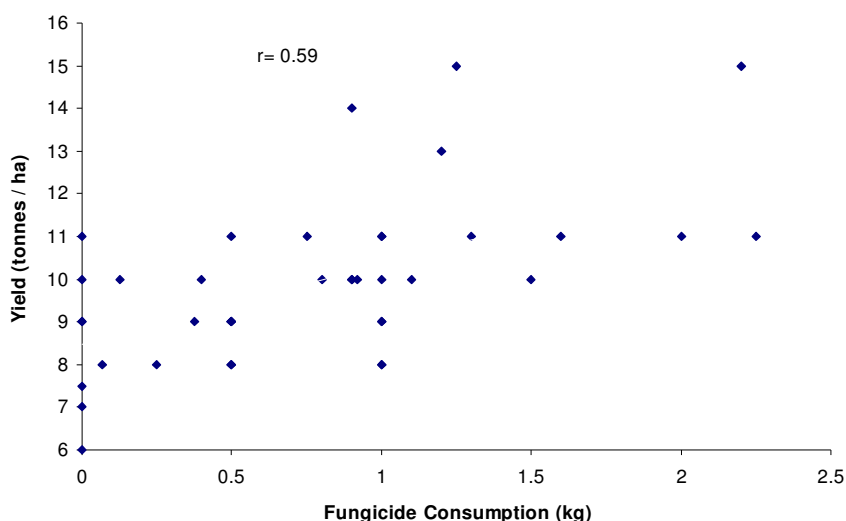
	Mean	Max	Min	SD	95% confidence interval	
					Lower	Upper
$\text{t ha}^{-1}$	9.9	15	6	1.94	9.29	10.53



**Figure 1.** Correlation between a) total farm area and yield, b) proportion of wheat area and yield, c) yield (tonnes /ha) and tractor hp/ha, d) yield (tonnes /ha) and N consumption, e) yield (tonnes /ha) and fungicide consumption.

**Continued...**

Continued of figure1.



**Figure1-e.** Correlation between yield (tonnes /ha) and fungicide consumption.

consumption was significantly correlated with yield ( $r = 0.59$ ). Maybe, and just maybe, fungi reduced yield more than other pests or they were more active on the farms with higher yield than on other farms. It was noticeable that fungicide consumption in wheat production was extremely low and its effect on yield must be taken into consideration.

### The Model Results

#### Multiple Linear Regression Model

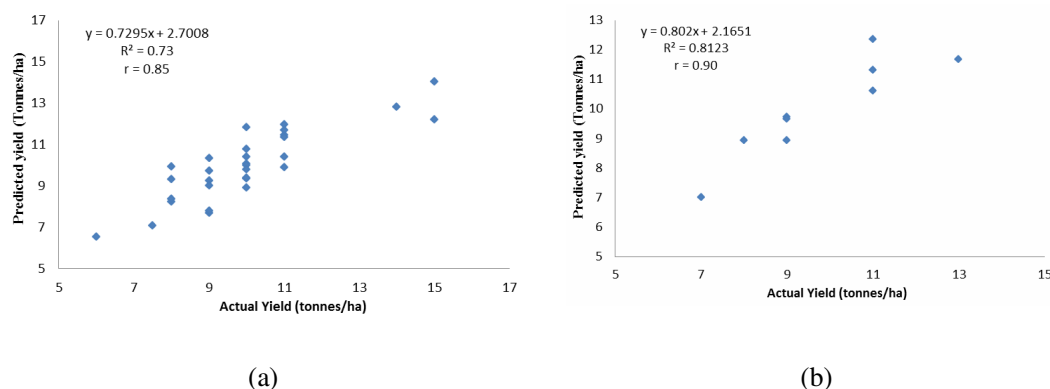
To predict wheat production, multiple linear regression and NN methodologies were developed. The multiple Linear Regression model has been extensively used in agricultural experiments evaluations, with positive expected linear effects and negative quadratic effects (Colwell, 1994). A simple model with the highest  $r^2$  is designed through a combination of forward, backward, and stepwise regression adjustments. Terms are maintained in the final model if they are significant at  $P = 0.05$

(Alvarez, 2009). In the first step, the relationship between wheat production and each input variable was tested with simple linear regression using the  $r^2$  as the decision criterion. Then, a multiple Linear Regression model was developed for predicting wheat production as follows:

$$Y = a_0 + a_1V_1 + a_2V_2 + \dots + a_nV_n + \epsilon \quad (3)$$

Where,  $a_0 - a_n$  are the regression coefficients,  $V_0 - V_n$  are the independent variables, and  $\epsilon$  is error.

The model was in linear form to represent linear relationships of dependent variable with the independent variable. For better comparison with NN model, 25% of samples were randomly selected for validation and 75% of samples were used for training. After running the model, predictions on validation data were estimated. A multiple Linear Regression model was fitted to wheat production data and accounted for around 81% of the variance in validation data. Figures 6 and 7 compare the predicted wheat production for training and validation data, respectively. The figures show that the correlations between actual wheat production and predicted yield for training



**Figure 6.** Relationships between observed and predicted wheat production a) training, b) validation, using multiple linear regression model.

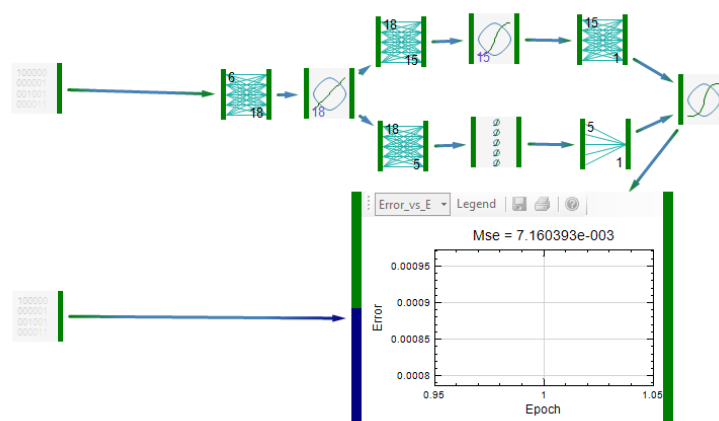
and validation are similar. The final *RMSE* for validation data was  $0.56 \text{ t ha}^{-1}$ .

### Neural Network Model

As stated previously, after testing different learning algorithms, neuron activation functions, and network structures, a modular network with two hidden layers were developed as shown in Figure 8. In the final model, Quick Prop learning method provided better performance than the other gradient decent methods and a linear function was selected for input layer. The

hyperbolic tangent ( $\tanh$ ) function was selected for the first part of hidden layer and logistic function was applied for the second part of hidden layer and output layer. The number of neurons in each layer was optimized using a genetic algorithm optimizer.

As shown in Figure 8, after input layer, the modular network is separated into two parts. The number of neurons in the first and second layers of the top part was optimized using a genetic algorithm optimizer that indicated 18 and 15 neurons for the first and second hidden layers, respectively. But, in the second part, the number of neurons was



**Figure 8.** Structure of modular network and number of neurons in each layer.

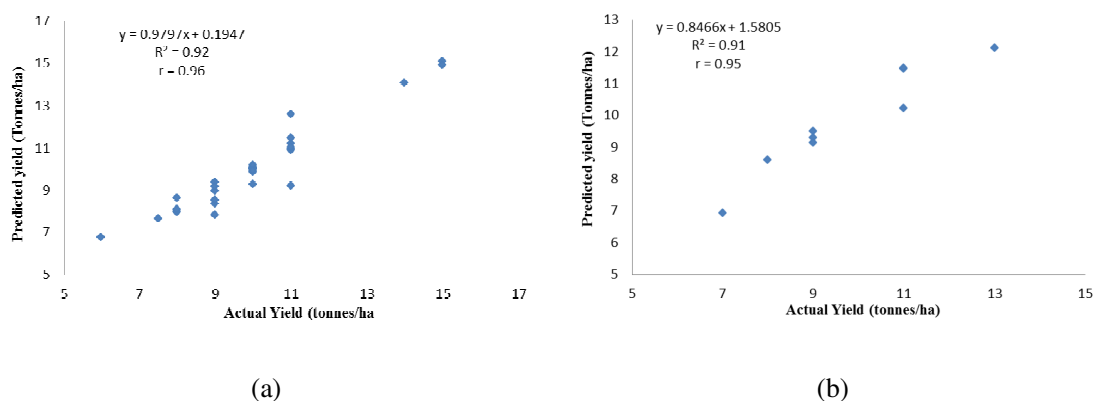


optimized to be 18 and 5 for the first and second hidden layers, respectively. The results are combined at the output layer to produce the final output, the wheat production.

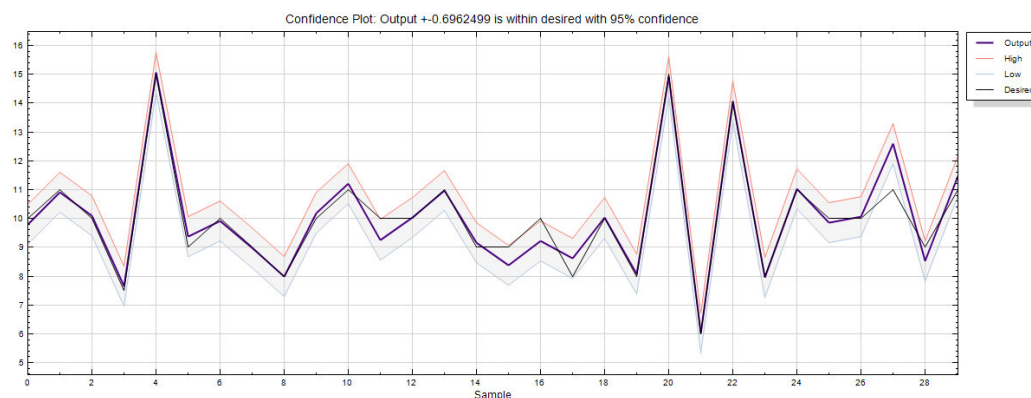
The NN model achieved the best results after 5,473 iteration, with scaled  $MSE = 7.2 \text{ E-3}$  (inputs and outputs were scaled between -1 and +1 for the neural networks model). The actual  $RMSE$  of the final NN model was estimated to be  $0.37 \text{ t ha}^{-1}$  on validation data. It was the lowest  $RMSE$  among several NN models examined in this study. Furthermore,  $RMSE$  of NN model on validation data was lower than that of the linear regression model. As shown in Figures 9 and 10, wheat production estimated by the NN accounted for 96 and 91% of the actual variability in

yield in training and validation data, respectively. Correlation between observed and predicted wheat production is very high with  $R^2 = 0.92$  and  $r = 0.95$  (training) and as the figures shows, it is higher than that of the multiple Linear Regression model. Results also showed that the correlation between actual and predicted wheat production in NN model ( $R^2 = 0.91$  and  $r = 0.95$ ) was much higher than that of Linear Regression model for validation data as well.

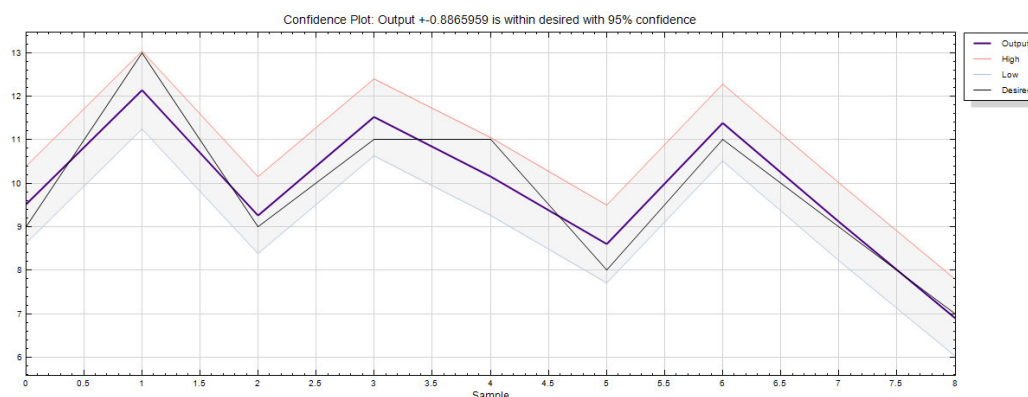
As shown in Figures 11 and 12, the final model predicted wheat production with an error margin of around  $\pm 0.70 \text{ t ha}^{-1}$  (training data) and  $\pm 0.89 \text{ t ha}^{-1}$  (validation data); and considering the uncertainties involved, this level of error is remarkable. In agricultural



**Figure 9.** Relationships between observed and predicted wheat production a) training b) validation, using artificial the neural networks model.



**Figure 11.** Predicted, observed, and the 95% confidence interval, for the wheat production based on the artificial neural networks model (training data).



**Figure 12.** Predicted, observed, and the 95% confidence interval, for the wheat production based on the artificial neural networks model (validation data).

production, there are several uncontrolled factors which could influence yield; therefore, the result of this study is very interesting and the final model can predict wheat production with acceptably small error.

The NN model can estimate wheat production and compare yield on farms easily. Farmers can use the model to explore the factors that have more potential to increase wheat production on their farms. Additionally, decision makers and scientists can estimate yield in different regions of Canterbury and they can evaluate the effects of different factors on wheat production.

Comparison between ANN model and multiple linear regression models showed that the correlation between the actual and predicted energy consumption for the ANN model was much higher than that for the linear regression model for both training and validation data; furthermore, *RMSE* of the ANN model on validation data was much lower than that of the linear regression model (Table 2).

## CONCLUSIONS

This study was the first time an ANN model was designed to predict agricultural production using direct factors as well as indirect factors, and it showed the potential of using indirect factors to predict agricultural production. The final neural network model, using a carefully selected set of six inputs from more than 140 different factors, can predict wheat production based on wheat area, irrigation frequency, machinery condition (tractor hp ha<sup>-1</sup> and number of passes of sprayer), and farm inputs (N and fungicides consumption) in Canterbury arable farms with an error margin of  $\pm 9\%$  ( $\pm 0.89$  t ha<sup>-1</sup>).

The results of this study showed the ability of ANN models to predict wheat production using heterogeneous data better than using a multiple regression model (as a common model used in agricultural studies), as shown in Table 2. Using dissimilar variables, such

**Table 2.** *MSE* and *RMSE* for training and validation of the multiple linear regression and ANN models.

	Linear		ANN	
	Training	Validation	Training	Validation
MSE	0.52	0.31	0.16	0.14
RMSE	0.72	0.56	0.40	0.38

as farm conditions and social factors would improve the ability of decision makers to look at the problem from different perspectives.

This study showed the ability of an ANN models to predict wheat production of farm inputs using indirect parameters. Improving the model to predict the yield of all farm products can provide more practical results for decision makers. It was clear that changing some of the effective variables in the short term was not possible; however, the model can help scientists and decision makers to find the best direction to increase farm production in the future.

Increasing the number of samples and testing more variables for at least five years can help design a model to predict the trend of agricultural production under different conditions. The results in this study may be considered as a first step in developing methods suitable for predicting wheat production for the whole Canterbury region by using social, technical, and geographical factors together. The method can be applied to other cropping areas of the world and to different crops.

## REFERENCES

1. Alvarez, R. 2009. Predicting Average Regional Yield and Production of Wheat in the Argentine Pampas by an Artificial Neural Network Approach. *European J. Agron.*, **30**: 70-77.
2. Breiman, A. and Graur, D. 1995. Wheat Evaluation. *Israel J. Plant Sci.*, **43**: 58-95.
3. Colwell, J. D. 1994. Estimating Fertilizer Requirements: A Quantitative Approach. CAB International, Wallingford.
4. Fang, Q., Hanna, M. A., Haque, E. and Spillman, C. K. 2000. Neural Network Modeling of Energy Requirements, for Size Reduction of Wheat. *ASAE*, **43(4)**: 947-952.
5. FAO. 2012. *FAO Statistical Yearbook 2012: World Food and Agriculture*. Food and Agriculture Organization of the United Nations, Rome.
6. Faramarzi, M., Yang, H., Schulin, R. and Abbaspour, K. C. 2010. Modeling Wheat Yield and Crop Water Productivity in Iran: Implications of Agricultural Water Management for Wheat Production. *Agric. Water Manag.*, **97**: 1861-1875.
7. Folberth, C., Yang, H., Wang, X. and Abbaspour, K. C. 2012. Impact of Input Data Resolution and Extent of Harvested Areas on Crop Yield Estimates in Large-scale Agricultural Modeling for Maize in the USA. *Ecol. Model.*, **235-236**: 8-18.
8. Gemtos, T. A., Markinos, A. and Nassiou, T. 2005. Cotton Lint Quality Spatial Variability and Correlation with Soil Properties and Yield. Precision agriculture'05. Papers presented at the 5th European Conference on Precision Agriculture, Uppsala, Sweden, pp. 361-368.
9. Hagan, M., Demuth, H. and Beale, M. 2002. *Neural Network Design*. PWS Publishing Company, Boston, USA.
10. Heinzow, T. and Tol, R. S. J. 2003. *Prediction of Crop Yields Across Four Climate Zones in Germany: An Artificial Neural Network Approach*. Centre for Marine and Climate Research, Hamburg University, Hamburg.
11. Hornik, K., Stinchcombe, M. and White, H. 1989. *Multilayer Feedforward Networks Are Universal Approximators*. Elsevier Science Ltd. Oxford, UK, **2**.
12. Jebaraj, S. and Iniyan, S. 2006. A Review of Energy Models. *Renew. Sust. Ener. Rev.*, **10**: 281-311.
13. Kaul, M., Hill, R. L. and Walthall, C. 2005. Artificial Neural Networks for Corn and Soybean Yield Prediction. *Agric. Sys.*, **85**: 1-18.
14. Khakural, B. R., Robert, P. C. and Huggins, D. R. 1999. Variability of Corn/Soybean Yield and Soil/Landscape Properties across a Southwestern Minnesota Landscape. *Precision Agriculture*, 573-579.
15. Kole, C., 2006. *Cereals and Millets*. Springer, New York, Berlin.
16. Kominakis, A. P., Abas, Z., Maltaris, I. and Rogdakis, E. 2002. A Preliminary Study of the Application of Artificial Neural Networks to Prediction of Milk Yield in Dairy Sheep. *Computers and Electronics in Agriculture*, **35**: 35-48.
17. Kronstad, W. E. 1998. Agricultural Development and Wheat Breeding in the 20<sup>th</sup> Century. *The 5<sup>th</sup> International Wheat Conference: Developments in Plant Breeding*, Kluwer Academic Publishers,



- Dordrecht, The Netherlands, Ankara, Turkey, , PP. 1-10.
18. Mitchell, D. O., Ingco, M. D. and Duncan, R. C. 1997. *The World Food Outlook*. Cambridge University Press, Cambridge, New York.
  19. Niska, H., Skon, J. -P., Packalen, P., Tokola, T., Maltamo, M. and Kolehmainen, M. 2010. Neural Networks for the Prediction of Species-specific Plot Volumes Using Airborne Laser Scanning and Aerial Photographs. *IEEE Trans. Geosci. Remote Sens.*, **48**: 1076-1085.
  20. Özdoğan, M. 2011. Modeling the Impacts of Climate Change on Wheat Yields in Northwestern Turkey. *Agric. Ecosys. Environ.*, **141**: 1-12.
  21. Ozkan, B., Akcaoz, H. and Fert, C. 2004. Energy Input-output Analysis in Turkish Agriculture. *Renew. Ener.*, **29**: 39-51.
  22. Papageorgiou, E. I., Aggelopoulou, K. D., Gemtos, T. A., Nanos and G. D. 2013. Yield Prediction in Apples Using Fuzzy Cognitive Map Learning Approach. *Computers and Electronics in Agriculture*, **91**: 19-29.
  23. Papageorgiou, E. I., Markinos, A. T. and Gemtos, T. A. 2011. Fuzzy Cognitive Map Based Approach for Predicting Yield in Cotton Crop Production as a Basis for Decision Support System in Precision Agriculture Application. *Appl. Soft Computing*, **11**: 3643-3657.
  24. Pimentel, D. and Pimentel, M. 2008. *Food, Energy, and Society*. CRC Press, Boca Raton, FL.
  25. Rosegrant, M. W., Agcaolli-Sombilla, A. and Perez, N. 1995. Global Food Projections to 2020: Implications for Investment 2020 Vision Discussion Papers 5. Intl Food Policy Res Inst.
  26. Safa, M. and Samarasinghe, S. 2011. Determination and Modelling of Energy Consumption in Wheat Production Using Neural Networks: A Case Study in Canterbury Province, New Zealand. *Ener.*, **36**: 5140-5147.
  27. Salehi, F., Lacroix, R. and Wade, K. M. 1998. Improving Dairy Yield Predictions through Combined Record Classifiers and Specialized Artificial Neural Networks. *Computers and Electronics in Agriculture*, **20**: 199-213.
  28. Samarasinghe, S. 2007. Neural Networks for Applied Sciences and Engineering: From Fundamentals to Complex Pattern Recognition. Auerbach, Boca Raton, FL.
  29. Schultz, A., Wieland, R. and Lutze, G. 2000. Neural Networks in Agroecological Modelling-stylish Application or Helpful Tool? *Computers and Electronics in Agriculture*, **29**: 73-97.
  30. Sharma, A. K., Sharma, R. K. and Kasana, H. S. 2007. Prediction of First Lactation 305-day Milk Yield in Karan Fries Dairy Cattle Using ANN Modeling. *Appl. Soft Computing*, **7**: 1112-1120.
  31. Singh, H., Singh, A. K., Kushwaha, H. L. and Singh, A. 2007. Energy Consumption Pattern of Wheat Production in India. *Ener.*, **32**: 1848-1854.
  32. Soares, J. D. R., Pasqual, M., Lacerda, W. S., Silva, S. O. and Donato, S. L. R. 2013. Utilization of Artificial Neural Networks in the Prediction of the Bunches' Weight in Banana Plants. *Scientia Horticulturae*, **155**: 24-29.
  33. Sözen, A. 2009. Future Projection of the Energy Dependency of Turkey Using Artificial Neural Network. *Ener. Policy*, **37**: 4827-4833.
  34. Statistics New Zealand. 1999. *A Regional Profile: New Zealand, Canterbury*. Statistics New Zealand, Wellington, New Zealand.
  35. Statistics New Zealand. 2004. *Water: Monetary Stock Report*. Statistics New Zealand, Wellington, NZ.
  36. Statistics New Zealand. 2008. *Agricultural Production Statistics (Final): June 2007*. Statistics New Zealand, Wellington.
  37. Statistics New Zealand. 2010. *Harvests Increase for Wheat, Barley, and Maize Grain*. Statistics New Zealand, Wellington, NZ.
  38. Sudduth, K. A., Drummond, S. T., Birrell, S. J. and Kitchen, N. R. 1997. Spatial Modeling of Crop Yield Using Soil and Topographic Data. Precision agriculture '97. Volume I. Spatial variability in soil and crop. Papers presented at the First European Conference on Precision Agriculture, Warwick University, UK, 7-10 September 1997.
  39. Uno, Y., Prasher, S. O., Lacroix, R., Goel, P. K., Karimi, Y., Viau, A. and Patel, R. M. 2005. Artificial Neural Networks to Predict Corn Yield from Compact Airborne Spectrographic Imager Data. *Computers and Electronics in Agriculture*, **47**: 149-161.

40. Wendroth, O., Jurschik, P., Giebel, A. and Nielsen, D. R. 1999. Spatial Statistical Analysis of On-site Crop Yield and Soil Observations for Site-specific Management.

Proceedings of the Fourth International Conference on Precision Agriculture, Pts a and B.

### پیش‌بینی تولید گندم با استفاده از شبکه های عصبی و مطالعه فاکتورهای غیر مستقیم: مطالعه موردی در استان کانتربری نیوزلند

م. صفا، س. ساماراسینگ، م. نجات

#### چکیده

یک روش شبکه مصنوعی عصبی (ANN) برای تولید نوعی از گندم مورد استفاده قرار گرفت که از بین داده‌هایی وسیع از ۴۰ مزرعه در کانتربری، نیوزیلند گردآوری شده است. میانگین تولید گندم ۹.۹ تن بر هکتار برآورد شده است. مدل نهایی توسعه (ANN) قادر به پیش‌بینی تولید گندم تحت شرایط مختلف و سیستم‌های کشاورزی با استفاده از عوامل مستقیم و غیر مستقیم است. پس از بررسی بیش از ۱۴۰ عامل مختلف ۶ عامل به عنوان عوامل اصلی و تاثیر گذار در داده‌ها انتخاب شد. مدل نهایی (ANN) پیش‌بینی تولید گندم را در کانتربری با خطای مجاز  $0.9 \pm$ ٪ ( $0.89 \pm$  تن بر هکتار) بر اساس این موارد انجام داد: شرایط مزرعه (منطقه کاشت و تعداد آبیاری)، وضعیت ماشین‌آلات (میزان اسب بخار تراکتور بر هکتار و تعداد گذر) و نهاده‌های مزرعه (میزان استفاده از کود نیتروژن و قارچ کشها).